Exploring Out-Patient Behaviors in Claim Database: A Case Study Using Association Rules

Yu-Chun Chen, Msc[+], Shiao-Chi Wu, PhD[+]
[+]Institute of Health Informatics and Decision Making,
National Yang-Ming University, Taiwan, R.O.C

Abstract

Patient behaviors are affected by so many factors that makes it is not easy to describe by simple measures. Our approach bases on finding association rules, which is widely used in marketing to identify their customers. We found rules describing complex patient behavior from huge claim dataset and we grouped them into several groups for a better understanding.

Background

Utilization of health resources is of interest from both medical and health policy perspectives. Several studies have found that most visits are made by a small number of patients who consume most of the resources available through the physicians[1]. A certain patients who are characterized by changing their medical sources frequently are called "doctor-shopping patients". "Doctor shopping" behavior has aroused great attention because there are lots issues behind this phenomenon. People tends to be shoppers because of various reasons such as simply their own personal characteristics[2], seeking for drugs[3], looking for second opinion[4], poor health care quality[5], moral hazard induced by health provider, and so on. Besides of several unavoidable interviewing error or bias introduced by traditional questionnaires, we look for interesting patterns among the complex behaviors recorded in the national wide health insurance data. We not only identify these patients but also outline the patients by profiles.

Method

Sourcing and preprocessing the data

Patients with annual visit to medical center counts greater than two were selected for our study. Three data models: Demographic data, Transactional data, and Relation data are aggregated from national health insurance claim database. All numerical attributes such as age, payment, annual visit count, annual hospital count…etc are discretized according to their distribution density. Records with poor quality are processed as well. Patients who had visited more than two hospital because of the same episode are classified as Shoppers.

Finding Emerging Patterns

All visits made by shoppers are separated from non-shoppers. We find all emerging patterns with support equals to 1% in shoppers' visits. Emerging patterns threshold are set to 2.5 as suggested by[6].

Evaluation of discovered rules

Mined rules are grouped as "rule cover" to ease interpretation[7]. We verified mined rules by logistic regression model.

Results

Among 50,000 patients, more than 10,000 "behavior rules" are mined. We successfully prune those rules into a manageable size by using post process technique.

Our exploratory shows that out-patients in Taiwan utilize their health service very subjectively. They would visit one department of hospital whereas another department of other hospital.

Some factors affect shopping behaviors such as age, sex, rate of utilization, chronic diseases, mental disorders and so on are compatible with existing studies. Emerging patterns show more complex conditions than traditional regression technique, those might provide interesting hidden insights of visiting behaviors.

Conclusions

Here we applied the standard data mining technique of association rules as a basis to explore patient visit data. We are able to see the real things in both detail and general levels. We think methods of this type could be useful in exploring in public health systems those fields concerned about complex behaviors. The techniques that we used here could readily be improved as new algorithms are developed recently.

References

1. von Korff, M., et al., *Disability and depression among high utilisers of health care.* Arch Gen Psychiatry, 1992. **49**: p. 91-100.
2. Bornstein, B., D. Marcus, and W. Cassidy, *Choosing a doctor: an exploratory study of factors influencing patients' choice of a primary care doctor.* J Eval Clin Prac, 2000. **6**(3): p. 255-62.
3. Longo, L., et al., *Addiction: part II. Identification and management of the drug-seeking patient.* Am Fam Physician, 2000. **61**(8): p. 2401-8.
4. Sato, T., et al., *Second opinion behavior among Japanese primary care patients.* Br J Gen Pract, 1999. **49**(444): p. 546-50.
5. Lo, A., et al., *Doctor-shopping in Hong Kong: implications for quality of care.* Int J Qual Health Care, 1994. **6**(4): p. 371-81.
6. Dong, G. and J. Li. *Efficient Mining of Emergin Patterns: Discovering Trends and Differences.* in *KDD 99*. 1999.
7. Toivonen, H., et al. *Pruning and grouping of discovered association rules.* in *ECML-95 Workshop on Statistics, Machine Learning, and Knowledge Discovery in Databases*. 1995. Heraklion, Greece.